DOCUMENT RESUME

ED 417 270 CE 075 689

AUTHOR Chandler, James A.; Fulwood, Robert H.; Hill, Lynn M.; Sharp, Jeanie B.; Vincent, Keith A.; Womac, Charlie H.; Campbell, Clifton P.

TITLE Vocational Course/Program Evaluation: A Four-Phase Approach.
INSTITUTION Tennessee Univ., Knoxville.
PUB DATE 1997-12-00
NOTE 53p.
PUB TYPE Guides - Non-Classroom (055)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Adult Education; *Course Evaluation; *Evaluation Methods; Job Training; Labor Force Development; Models; Postsecondary Education; Program Evaluation; Records (Forms); Secondary Education; *Systems Approach; *Vocational Education

ABSTRACT

A four-phase systematic approach to evaluating human resource development (HRD) programs and/or courses was developed. The approach consists of four distinct, yet interrelated phases: select evaluation targets; develop the evaluation plan; collect and analyze information/data; and prepare an evaluation report. First, 23 potential evaluation targets in the following 6 categories were identified: course/program development; course/program delivery; trainee achievement; instructor performance; graduate performance in the workplace; and cost-benefit analysis. Next, the following steps in the evaluation-planning phase were detailed: specify the evaluation's purposes and objectives; identify sources of information; identify information/data collection methods; and prepare a program evaluation activities schedule. Available types of instruments for collecting qualitative and quantitative data were discussed along with steps for developing a balanced approach to evaluation by maintaining a constant awareness of the company's vision and working toward that vision. A detailed plan for evaluation reports was outlined that included guidelines for presenting evaluation findings and developing an action plan based on the results. (The paper contains 19 references. Appended are the following: biographical information about the authors; chart illustrating the concept of a balanced evaluation scoreboard; chart for recording a performance measurement index design group; and glossary.) (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 417 270

Vocational Course/Program Evaluation: A Four-Phase Approach

James A. Chandler
Robert H. Fulwood
Lynn M. Hill
Jeanie B. Sharp
Keith A. Vincent
Charlie H. Womac
Clifton P. Campbell

The University of Tennessee, Knoxville

December, 1997

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
J Chandler
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

CE 075 685

BEST COPY AVAILABLE

Vocational Course/Program Evaluation:

A Four-Phase Approach

Introduction

Fayol, in his classical organizational theory, considered a well-managed organization to be one that is following a controlled path leading to organizational goals. Managers are responsible for keeping the organization on that path. In order to do this, management needs information resulting from "measuring performance against standards, detecting if and [determining why] targeted performance" was not achieved, and taking action to correct any problems discovered (Powers, 1995).

In the context of human resource development (specifically, occupational training), maintaining controlled progress toward organizational goals is at the very heart of vocational course/program evaluation (hereafter referred to as "evaluation"). The purpose of evaluation is to provide valid, meaningful information that will allow "management by exception." That is to say, if the evaluation indicates that the course/program is achieving the desired results, there is no need for any corrective action. On the other hand, if a discrepancy exists, a well-designed and conducted evaluation can assist by identifying the discrepancy, as well as by providing corrective information.

In the following pages, the authors present a systematic four-phase approach to evaluation. Each phase is distinct, yet all are interrelated. Phase 1 (Select Evaluation Targets) supplies a series of recommendations for identifying evaluation targets that should be included in the evaluation. Phase 2 (Develop the Evaluation Plan) identifies and discusses a series of steps to follow in planning the evaluation. Phase 3 (Collect and Analyze Information/Data) outlines a process for collecting and analyzing data during the evaluation. Phase 4 (Prepare an Evaluation

Report) discusses how to use data analysis results to draw conclusions, develop recommendations, and prepare the final evaluation report for the decision makers.

Phase 1: Select Evaluation Targets

Phase 1 focuses on selecting what is to be measured and how to measure it. These choices should be governed by the needs of decision makers involved in both training program design and management of the organization for which the training and evaluation are being performed. "The primary role of program evaluation is to provide relevant information to key decision makers so they can make informed judgements" (Rhoades, 1991, p. 64).

Based on these needs and considering the specific purpose(s) and objectives of the training, the purpose(s) of the evaluation must be defined. L'Angelle (1996) states "Evaluation must measure (a) how well the program works . . . and (b) the quality of the finished products" (p. 223). In other words, the evaluation must be both formative and summative. Gunn (1987) states that political concerns, unfortunately, are often part of the actual purpose of evaluations. When this is the case, the evaluator should be aware of it as the evaluation purposes are formulated.

Based on the above information, the evaluator must select evaluation targets. In this report, "evaluation targets" refers to processes or outcomes one measures in order to guage the effectiveness of the training program. The results of these measurements should furnish decision makers with information they need to make wise choices about the fate of the training program. For example, evaluation could be undertaken to help a manager make a decision about the future status of a program (i.e., whether it should be continued, expanded, reduced, modified, or

discontinued). On the other hand, the same evaluation might help a training specialist gain insights concerning how to improve some aspect of the training process.

Table 1 lists potential applications of evaluation results. Items 1.1 through 1.4 would be useful primarily for those involved in planning and implementing the training process. Also, these items would be used primarily in formative evaluation. Items 1.5 and 1.6 would be most useful for managers in the organization where training and evaluation was or is being conducted (hereafter referred to as the "client organization") and would represent summative evaluation.

Under each major item in the table is a list of potential evaluation targets which could be measured to assess effectiveness of respective major items. For example, the effectiveness of the training course/program could be assessed by measuring the percentage of training program graduates who obtained jobs related to the training. If the percentage is low, training planners probably did not accurately assess the projected demand for skills being taught in the training program. Such projections should be made early in the training planning stage and should have a major influence on selection of skills to be taught in the training program.

It is important that evaluation targets selected furnish information which meets the needs of decision makers in both the training program and the client organization. Choices of evaluation targets made during the evaluation planning stage will have major impacts on the effectiveness of the evaluation program. Selecting targets which address the needs of both groups of decision makers, helps curriculum planners work effectively with those who are in key decision-making roles. The training program decision maker needs targets that furnish information useful in improving the training program. The decision maker in the client organization needs targets which help assess effectiveness from a cost-benefit perspective. Also,

Table 1

Potential Evaluation Targets

1.1 **Course/program development**
Labor market demand (need for graduates) and supply (number of skilled workers available and in training)
Needs assessment
Job and task analysis
Instructional content
Instructional materials
Instructional setting

1.2 **Course/program delivery**
Learning objective relevance/accuracy
Instructional method selection and use

1.3 **Trainee achievement**
Selection criteria (admission to program/course)
Test performance
Skill proficiency
Behavioral change

1.4 **Instructor performance**
Plan instruction
Conduct instruction
Evaluate training

1.5 **Graduate performance in the workplace**
Application of acquired knowledge and skills on the job
Work attitude/ethic
Employment record
Career progression

1.6 **Cost/benefit analysis**
Analysis cost
Development cost
Delivery cost
Evaluation cost

Note. Adapted from "An Approach to Program Evaluation," by D. D. L'Angelle, 1996, in C. P. Campbell (Ed.), Education and Training for Work: Planning Programs. Lancaster, PA: Technomic Publishing.

targets must be compatible with the structure and culture of the organization. Some individuals whose needs should be considered in selecting evaluation targets are training managers, instructors, instruction designers, management decision makers, advisory committees, trainees, and human resource development (HRD) career counselors. The needs of each of these in relation to selecting evaluation targets are discussed below.

Training Managers

Training managers perform various roles including (a) manager of the organizational learning system; (b) operational manager of planning, organizing, staffing, controlling, and coordinating the training department; (c) strategist responsible for long-term planning; and (d) marketing specialist responsible for advancement of training within the organization.

The evaluation information this group needs often differs depending on the role(s) being performed at the time. Managers also use evaluation information in decisions relating to personnel practices, financial accountability, marketing, and when deciding the fate of training programs. Further discussion of the needs of management decision makers is presented later.

Instructors

An evaluation system can be designed to identify deficiencies and strengths of instructors, helping them improve their performance. Evaluation targets should relate to effectiveness of the instructor and increase awareness of strengths and weaknesses. These strengths and weaknesses would relate to both the presentation style of the instructor and how effectively visual aids and other instructional media are used.

Instruction Designers

The instruction designer designs, develops, and often evaluates courses and programs, but seldom implements them.

> At the heart of all learning programs and training activities is their design, a blueprint from which to construct all learning in the organization. Without a properly designed program, learning will not be consistent, nor will desired results become evident (Gilley & Eggland, 1989, p. 22).

Instruction designers have information needs in designing or altering training programs to ensure proper material content and presentation methods appropriate to trainees' abilities and prior knowledge and skill.

Consultants

"Consultants range from technical experts for instruction and program design to facilitators who solve problems to bring about change. Consultants may be internal or external to the organization" (Gilley & Eggland, 1989, p. 21). Regardless of the consultant's role, he will require specific evaluation information to determine program effectiveness. A consultant's need for information can be triggered by the discovery of a performance problem, a request for certain training from management, a general observation of a problem, a need to improve productivity, or a requirement to comply with government regulations.

Management Decision Makers

Management decision makers must have valid and reliable information on which to base decisions. Evaluation findings can support decisions making them defensible and justifiable to the trainee, staff, administration, and the public. Management decision makers may also need

information for use as a financial accountability aid. Managers need to know if performance improvement or other benefits resulting from training justify the cost of the training program. To determine this, evaluation targets must accurately reflect training program benefits.

Advisory Committees

Advisory committee members may include representatives from industrial councils, parent groups, Chambers of Commerce, employment services, boards of trustees, and so forth. Advisory committee members' opinions have potential for contributing to the evaluation, regardless of their chain of command level. Information needs of advisory committees vary depending on the type of relationship that the committee has with the organization.

Trainees

Trainee characteristics impact curriculum content. Consequently, it is important to determine their (a) general and applied skills, (b) interests, (c) motivators, and (d) maturity levels. The use of pre-testing can help determine these characteristics.

Trainees are also involved in the evaluation process as decision makers. Follow-up studies of graduates provide current trainees with data about former trainees with similar or comparable experiences and their success in the workforce after graduation. The trainee can use this information as a guide for selecting a program of study. Such information could also be useful as a marketing aid for prospective trainees.

HRD Career Counselors

Counselors normally require information about trainee achievement, abilities, strengths and weaknesses, and concerns about the training program. Insights gained from this information helps them improve trainee performance.

Phase 2: Plan the Evaluation

Before beginning to plan an evaluation, the evaluator must identify its purpose, what is to be evaluated, and who is to be involved. If the evaluation process will involve a team, team members should begin to carefully identify the steps to be included in the process. The team must exercise care to develop an evaluation procedure that will assist the decision makers (i.e., the evaluation audience) in making the proper choices and reaching the correct conclusions. In order to achieve this mission, the evaluation must be credible. Kenneke (1987) stated that credible evaluations: "are perceived as accurate . . . [the] findings [are] discussed and used by the audience. What people accept as credible information is influenced by the evaluation design" (p. 2).

In addition to being credible, the evaluation must provide relevant, useful information that contributes to the audience's decision-making process. Therefore, the findings of the evaluation must be meaningful, specifically-stated, and quantitatively-measurable. Excessive, confusing and/or meaningless facts and figures must be excluded.

In the following pages, the authors present a series of step-by-step recommendations and/or instructions which provide assistance in developing and conducting an evaluation of a vocational course or program.

Step 1 - Specify the Evaluation Purposes and Objectives

As the evaluation plan is being formulated, the primary concern is that the evaluator (or team) focuses on the specific purposes and the objectives that are to be achieved. That is to say, the exact purpose of the evaluation process should be stated clearly and precisely. In addition, the evaluator (or team) should develop policies, procedures, and guidelines to govern the

evaluation process that clearly define the mission of the evaluation, and that are consistent with the mission of the institution. The objectives, conditions, policies and procedures, etc. must be written in a style that is easily understood by everyone involved.

The next concern is how the instructional information is delivered to trainees. For example, the amount of knowledge retained by trainees during lecture or discussion sessions can (and will) vary greatly from the amount that will be retained during laboratory or practicum sessions. In the former situation, learning depends on lecture, discussion, printed matter, and audiovisual and/or other instructional aids. The latter situation, requires the trainees to demonstrate attainment of a skill or other competence related to the training provided by actually using the skill to perform a task.

Finally, in most instances, the evaluation planning step should include some form of cost-benefit analysis. This procedure allows the evaluator(s) to establish a time schedule and/or financial commitment for achieving the organization's goals and objectives. Cost-benefit analysis may not be relevant or appropriate following compliance training, new product assembly or production training, and training due to organizational change or realignment. The need for, as well as the type of, cost-benefit analysis should be determined for each evaluation.

To summarize this step, planning for training program evaluation requires that the following information be included when specifying the purposes and objectives of the evaluation process: "(a) Who wants the evaluation? (b) What do they want to know? (c) What design will provide credible information? (d) What should be measured? (e) How should information be reported?" (Kenneke, 1987, p. 6).

In an unpublished manuscript, Campbell (1996) reported that several training specialists and human resource development managers identified the following eight potential evaluation purposes as being the most fundamental:

1. Determine the effectiveness and efficiency of the various course/program components: instructional content, materials, methods, media, setting; testing/evaluation; schedule; instruction; operation; management; etc.
2. Determine what, if any, part of the training needs improvement
3. Identify course/program graduates' reactions to the training they completed
4. Determine if the course/program should be continued, expanded, reduced, modified, or discontinued, based on labor market demand
5. Broaden occupational offerings according to trainees' needs
6. Determine the effectiveness of the training in terms of the application of knowledge and skills on the job, graduate achievement, and the value of the knowledge and skills acquired
7. Meet accountability demands
8. Meet requirements imposed by licensing and accreditation, if appropriate (p. 6)

In order to provide the reader with a partial model, a sample evaluation plan is provided in Table 2.

Step 2 - Identify Sources of Information

When designing a training program evaluation, it is important to know who wants to use or benefit from the results. Generally speaking, audience members are interested in all types of information regarding training programs. For instance, they may want to know (a) how the program was designed, (b) whether trainees are being provided with relevant information,

Table 2

Partial Evaluation Plan

Evaluation purposes:

1.0 Determine if the program should be continued, expanded, reduced, modified, or discontinued, based on labor market demand (need for graduates) and graduate placement.

2.0 Determine if the course/program was effective in terms of (a) the relationship of instructional content to workplace needs, (b) instructor performance, (c) trainee achievement, and (d) the application of knowledge and skills on the job.

Evaluation objectives for evaluation purposes 1.0 and 2.0: Is there evidence that -

1.1 Graduates are working in the job they were trained for?

2.1 Instructional content is consistent with job requirements from the perspective of the Advisory Committee?

A. Was a job and task analysis conducted in order to identify the required knowledge and skills?

B. Are knowledge and skill requirements traceable to the learning objectives and to criterion-referenced tests?

2.2 The instructional materials are effective in helping trainees learn and retain information?

2.3 Proven methods of instruction are used to help trainees learn?

2.4 The program is operating effectively by meeting job performance requirements?

2.5 A criterion-referenced testing program, including written and performance tests, is used to measure trainee attainment of learning objectives?

Note. Adapted from "An Approach to Program Evaluation," by D. D. L'Angelle, 1996, in C. P. Campbell (Ed.), Education and Training for Work: Planning Programs. Lancaster, PA: Technomic Publishing.

(c) whether learning objectives are being attained, (d) whether the teaching process is effective, (e) whether trainees are mastering job skills, and (f) whether program graduates get jobs.

In order to obtain this information, as well as any additional information that may be needed, it is important to identify the information sources.

1. Administrative personnel can provide information to establish goals, direct activities, evaluate instructional performance, and implement organizational change.

2. Instructional personnel plan and conduct training activities and gather evaluation data.

3. Trainees demonstrate that the competencies gained during training resulted in increased productivity on the job.

4. Supervisors attest to trainees' performance improvement after training.

Step 3 - Identify Information/Data Collection Methods

This step deals with the design of the evaluation process. According to Kenneke (1987), "a good design produces believable information because it provides comparative data. This allows results from the program under evaluation to be judged against another program or standard . . . these conclusions are more defensible than ones that merely report results without benefit of a comparison group or standard" (p. 8). Three different methods are commonly used to identify and collect data for the evaluation of vocational training programs: control group design, non-equivalent design, and informal comparison design.

Control group design. Trainees are randomly assigned to one of two groups. One group is trained using a new method, material, or program. This group is called the experimental (E) group. The second group of trainees, called the control (C) group, does not receive the training. Trainees in both groups take a pre-test to ensure that all trainees are at about the same skill or

knowledge level. However, if any of the trainees' pre-test scores are above a predetermined level, those trainees are excused from further participation in the study. After the experimental group has completed the training program, a post-test is administered to both groups. If the average (mean) post-test score of the E-group members is significantly higher than that of the C-group trainees, the difference between their respective scores can be attributed to the training program.

According to Kenneke (1987), there are variations on this design. The true control group approach may be altered so both groups are measured only at the completion of the treatment. This variation is appropriate when the pretest might bias program effects (when measuring attitudes) or when a suitable pre-test is not available (p. 10).

Non-equivalent design. This is a variation of the original pre-test-post-test design. Trainees are organized into nonrandom groups. This procedure can be beneficial when trainees cannot be grouped randomly because they are affected by the location, times, or other factors of the training program. Still another variation periodically tests experimental and non-equivalent groups of trainees. Both groups are tested before, during, and after the training program is implemented.

> Control group designs require assignment of trainees, classrooms or workshops to programs; comparisons of groups and rigorous control of variables. These controls enable one to say that evaluation results are due to the treatment, and not to other variables. True control group designs are highly credible means of determining whether or not a program is effective. These designs indicate effectiveness: however, they do not tell one what deficiencies exist, or how to correct them (Kenneke, 1987, p. 10).

Informal comparison design. This design does not collect information using any formal means. Data are gathered from alternate sources, such as employee personnel records, training program records, etc. Collectively, these data are used to form and document conclusions regarding the effectiveness of the training. For example, an evaluator may want to determine whether the production capacity of a particular machine is adequate enough to increase the level of production. To find the answer, the evaluator would analyze the production levels of all equipment and compare the current levels with the desired standards. In training, this is called a "needs analysis." In other words, if a training program evaluation shows that post-training performance is not consistent with expectations or standards, a discrepancy exists that must be corrected.

Step 4 - Prepare a Program Evaluation Activities Schedule

The purpose of the Program Evaluation Activities Schedule is to provide a written plan of events that are expected to occur during each stage of the project. It lists the sequenced activities, person(s) responsible, and a projected time frame during which each activity is to be completed. In developing an evaluation plan, one member from the team is selected to manage the process and ensure that each activity is completed by its end date. Table 3 provides a sample schedule.

Since good evaluations are carefully planned and executed, the merit of the evaluations are directly related to the amount and type of planning that has gone into them. Rhoades (1991) outlined the following six steps to guide the evaluation process:

Table 3

Program Evaluation Activities Schedule

Title of Training Program: _____ Evaluation Coordinator: _____

Activities	Person(s) responsible	Start & due dates	Work days
	Planning		
1. Form evaluation team	X. Y. Zee	3/01 - 3/02	2
2. Specify purpose and objectives of evaluation	A. B. Cole	3/04 - 3/05	2
3. Select evaluation tasks	A. B. Cole	3/06 - 3/08	3
4. Assign task responsibilities	A. B. Cole	3/09 - 3/10	2
	Before Training		
1. Develop evaluation plan	D. E. Franks	3/08 - 3/12	5
2. Present to key decision makers	A. B. Cole & D. E. Franks	3/15 - 3/15	1
3. Develop evaluation questions	G. H. Ioni	3/20 - 3/24	5
4. Develop cost-benefit analysis	G. H. Ioni	3/20 - 3/24	5
5. Print survey forms	D. E. Franks	3/27 - 3/28	2
	During Training		
1. Distribute survey to trainees after program completion	J. K. Lewis	5/01 - 5/01	1
	After Training		
1. Analyze evaluation data	J. K. Lewis	5/02 - 5/04	3
2. Formulate recommendations	J. K. Lewis	5/06 - 5/10	5
3. Write report	G. H. Ioni & J. K. Lewis	5/04 - 5/11	8
4. Present report findings	J. K. Lewis	5/12 - 5/12	1
5. Develop action plan	A. B. Cole	5/16 - 5/20	5

Note. Adapted from Rhoades, J. (1991). Evaluate regular and cooperative vocational programs: Module. Climax Springs, MD: Quality Training Specialists. (ERIC Document Reproduction Service No. ED 356-337)

Total time: 50 Days

1. Determine who should be involved. These could be the decision makers and/or sources of information which were identified in Phase 1 and Phase 2, respectively, who could provide additional information and clarify questions as needed.

2. Identify activity tasks. Each task should be explained in detail to the evaluation group in order to avoid unnecessary delays in meeting the projected due dates.

3. Assign personnel responsibilities. One person responsible for the overall activities schedule, an evaluation director, for example, should be selected among the evaluation group to assume this leadership role. Next, one person, or if a team effort, a group leader, should be selected from the remaining group members and has the responsibility to ensure that the assigned tasks are completed on time. Depending on the level of experience required for the project, this person can volunteer for the position, be nominated by the group, or be appointed by the evaluation director.

4. Sequence the tasks. List all the tasks and identify those which are dependent upon another task. Put the tasks in the order in which they must be completed. Many activities can be conducted simultaneously.

5. Set target dates. A projection time frame for the beginning and due dates for each task should be discussed and included in the activities schedule. This not only reinforces accountability, but it also provides a visual time line for the complete evaluation process.

6. Inform personnel. This should be an ongoing activity. If all members of the evaluation team are properly informed, they will be in a position to help one another, adequately explain the evaluation goals to others, make good decisions, and comprehend problems that are experienced by other team members. The tasks listed in Table 3 are not exhaustive, but serve as

an example to demonstrate how items could be sequenced during each stage of the evaluation process. This table provides a summary of the planning process and remains consistent with the purpose of this training manual — to provide a model that illustrates the evaluation process, a framework that can be modified and adapted based on the needs and expectations of the decision makers.

Phase 3: Collect and Analyze Information/Data

Questions such as the following were asked in Phase 2: What is to be evaluated? Who is to be involved? What is the purpose of the evaluation process?

Similar questions are asked in Phase 3, but from a different perspective. While Phase 2 was a general discussion of how to develop the evaluation plan, Phase 3 narrows the focus to consider the mechanics of collecting and analyzing information and data. This is followed by a discussion of how data are collected and analyzed in various types of training program evaluations: reaction surveys, follow-up evaluations, and impact surveys. The final section discusses evaluation strategies.

Capturing and Analyzing Relevant Data

In order for the results of a training evaluation to be useful, data collection methods must be carefully selected and implemented so that the resulting analyses are valid, reliable, and relevant to the needs of the evaluation audience. Since the instrument is the primary device for capturing data, it will be the principle focus of this section.

Phillips (1983) presents a series of questions which help guide planning and implementation of a training program evaluation. Four of these questions are listed below as

titles of subheadings in the discussion of how to capture relevant data and analyze it in a meaningful way (pp. 64-65).

Who will use the information [derived from data]? Data collection, analyses, and reporting should be planned to meet the needs of the audience by proper design of the data collection instrument. In designing the instrument and planning the analyses, the general educational background of the audience, including their familiarity with statistics should also be considered. This should result in a message which is understandable and useful for the audience. Some additional suggestions in presenting data and analyses are (a) design tables, graphs, and other means of presenting raw data to maximize communication, (b) limit data and information to what is needed by the audience, and (c) interpret analyses from the audience's perspective.

What facts are needed? The evaluator should provide facts relevant to the purposes of the evaluation and needed for the audience to make necessary decisions. Facts may relate to specific costs, output, time required to perform a task, quality of products, reactions to training, observations about training, and/or attitudes of graduates.

Based on insights gained by pursuing the information described above, the evaluator can select the type(s) of instrument(s) to be used in obtaining data. Some of the more commonly used types of instruments are (a) questionnaires, (b) attitude surveys, (c) tests, (d) interviews, and (e) performance records. An important consideration when developing or selecting a data collection instrument is its validity and reliability. Refer to the glossary (Appendix D) for definitions of these terms.

BEST COPY AVAILABLE

Questionnaires. Questionnaires are the most commonly used and versatile of the instrument types listed above. Some precautions in developing and reviewing questionnaires are (a) carefully word questions so the reader understands what information is requested, and (b) consider the questions discussed above as you design the questionnaire.

A questionnaire may contain questions asking the reader to (a) select from multiple alternative answers, (b) rank a list of possible choices, (c) provide an answer in the respondent's own words, and/or (d) check all responses they feel should apply.

Attitude surveys. One use of attitude surveys is to detect the effect of training on the attitudes of participants. When used for this purpose, attitude surveys are usually administered both before and after training, and the results of the two surveys are compared for each subheadings in the discussion of how to design or select an instrument to capture relevant data and analyze it in a meaningful way.

How will the data be used? By reviewing the purpose(s) of training evaluation and training evaluation audience needs, the evaluator should be able to collect data in a form which facilitates progress toward the goals of evaluation. Data can be used for a variety of purposes including (a) measuring effects of training on the graduates' attitudes toward work, (b) determining effects of behaviors learned during training on progress toward organizational goals, (c) determining effects of behaviors learned during training on progress toward the graduates' personal job performance goals, (d) comparing results of training to the objectives of the training program, (e) determining the extent of transfer of trained behaviors to the work environment, (f) defining the most effective and most ineffective components of the training

program, (g) performing a cost/benefit analysis, and/or (h) identifying graduates who gained the most and those who gained the least from the training.

How will the data be analyzed? While designing instruments for capturing data, the evaluator should also consider statistical analyses most appropriate for addressing evaluation purposes. This is also the time to consider how to present the data, results of data analyses, and interpretations to managers, training graduates, and others. If this question is not considered, meaningful statistical analyses and other manipulations of data may not be possible at the close of the evaluation process, thus rendering the evaluation effort and expense feudal.

Attitudes are so complex and are affected by so many variables, it is not likely an attitude survey alone will adequately assess employee attitudes. Other sources of information such as observations by supervisors, co-workers, and others as well as interviews may be used in conjunction with attitude surveys to get more useful information. The evaluator may save considerable time and expense by purchasing an existing attitude survey instead of developing one. However, the evaluation should be certain the survey instrument will furnish the information needs.

Tests. Tests are useful in quantifying effectiveness of training by comparing the results of the same test given to the same group of trainees before and after training. The difference in scores on the two tests, referred to as a "gain score," is assumed to reflect the amount trainees learned during training. There are several different types of tests. Some of the more common types are norm-referenced tests, criterion-referenced tests, and performance tests.

In the norm-referenced test, the group's or individual's score is compared to the score of another similar group or individual. This comparison standard or norm score may be the average

score of the group or a similar arrangement in which scores are evaluated based on a score or scores representing the norm. Another example of a norm-referenced test is when all members of the group taking a test are ranked. This ranking is norm referenced; it is referenced to the scores of the remainder of the group.

The criterion-referenced test measures performance relative to standards established based on task analysis of the job the training program is helping the trainee learn. The trainee is evaluated based on definite scores being above a specified minimum, not according to score in relation to scores of other group members.

The performance test evaluates trainees based on how well they demonstrate mastery of a trained skill, knowledge, or attitude. Performance tests are commonly used for on-the-job training and vocational training.

Interviews. A training program evaluator can often obtain information by interviewing trainees, supervisors, or other interviewees that they would be reluctant to share in a questionnaire or other written data collection instrument. However, disadvantages of the interview method are that (a) it requires more time to administer; and (b) the personality of the interviewer, the interview setting, or other factors may affect the interviewee's responses.

Performance Records. The evaluator may obtain data related to training evaluation objectives from existing data in personnel files, company financial records, and other written sources. Data collection from records should be planned so that only data relevant to evaluation of the specific training program is obtained.

The types of instruments described above can be used for a number of training program evaluation applications. Some of these applications are described in the following sections.

Reactions To Training

Measurement of customer satisfaction, or trainee satisfaction, can be used to determine if a training course or program has been successful (Willyerd, 1997, p. 53). What trainers call "smile sheets" are usually associated with Kirkpatrick's 1st level of evaluation of training (Phillips, 1991). This method is too subjective to provide a reliable report of how well the training objectives were met, whether applied to the program as a whole or to individual items such training materials, media, methods, setting, etc.

To improve exactness or "rigor" of evaluation instruments, the following list could provide guidance.

1. Limit questions to those which trainees are qualified to answer.
2. Select appropriate evaluation criteria such as, learning objectives, instructional media, quality of instruction, materials, trainee involvement, etc.
3. Keep statements and questions clear and easy to understand.
4. Select a proper rating scale.
5. Provide blank lines for additional written responses.
6. Collect feedback at timely intervals, such as during training, at the end of training, or a few weeks after completion of training.

Follow-up Evaluation

Careful planning and design based on a pre-training needs analysis is essential because the employee must first learn skills and knowledge before applying them within the job environment. However, it is often erroneous to assume that trainees who made excellent gain scores and seem to have mastered knowledge and skills during training will automatically apply

them within the work environment. Often, end-of-training performance is not an accurate indicator of the degree to which graduates will apply trained behaviors when they return to the work environment (hereafter referred to as "transfer"). This transfer is the ultimate goal of training because trained behaviors will not benefit the organization or the graduate unless they are applied within the work environment. Thus, the ultimate goal of training evaluation is to determine if and to what extent transfer is occurring.

Before focusing on the follow-up evaluation process, it may be helpful to consider how follow-up evaluation fits into the overall training program evaluation process. Kirkpatrick (1994) proposed a four-level model for training program evaluation. Levels 1 and 2 he labeled "reaction" and "learning". "Reaction" represents the level of personal satisfaction the graduate experienced from the training. Did he feel it was presented well? Was the training site comfortable? The "learning" level refers to short-term effects of training on the graduate's attitudes, knowledge, and skills. Levels 1 and 2 would be assessed by tests given at the close of training. These first two levels were discussed in the previous section.

Levels 3 and 4 of Kirkpatrick's model are "behavior" and "results". "Behavior" refers to the lasting effect the training has on the graduate's job performance after he has been back on the job for a few months. "Results" is the level of evaluation which measures effects of the training in helping or hindering progress toward organizational goals. Level 3 is assessed by a follow-up evaluation several months (usually 3 or 4) after completion of training. Level 3 represents the extent of transfer. This section will explain the purpose of the follow-up evaluation and data collection and analysis methods used in measuring it. Measurement of Level 4 will be discussed in a later section entitled "Evaluation Strategies."

The Purpose of Follow-up Evaluation. The purpose of follow-up evaluation is to assess the extent to which the training graduate is habitually applying knowledge, skills, and attitudes taught in the training program (hereafter referred to as "trained behaviors") while performing his regular duties within the work environment. This transfer is a major goal of training since the organization commissioning training is not likely to benefit if the trained behaviors are not manifested within the work environment.

Data Analysis. A follow-up evaluation reflects a successful training program when graduates demonstrate proficiency in skills learned, knowledge gained, and attitudes conveyed by the training program. The training program is considered even more successful if graduates then habitually apply the trained behaviors in response to appropriate "triggering events or situations" (Zenke & Gunkler, 1985) within the work environment.

Data Collection Instruments Used. The most common data collection instrument used in follow-up evaluation is the questionnaire. More information on questionnaires was presented in a previous section within Phase 3. Questionnaires may be used to obtain information from the graduates, their supervisors, co-workers, customers, and others.

A second technique commonly included in the follow-up evaluation is the interview. Interviews may be used to obtain information from a variety of persons knowledgeable about the job performance of the graduate. However, the expense and time required to conduct interviews makes them more difficult to administer than questionnaires.

Tests and performance records are two other possible sources of data during the follow-up evaluation. The test is usually more expensive than either the interview or the questionnaire because it generally requires time to administer and evaluate and may require setting up

equipment and simulations. However, tests are especially useful in assessing skill proficiency and knowledge retention because they require the trainee to actually perform the task or furnish the information conveyed during training. Performance records are a good source of objective and quantifiable data. They are usually easily accessed and typically do not require selecting a mutually acceptable time for two or more people to meet as tests and interviews often do.

Data analysis for the follow-up evaluation generally consists of comparison tests and tests of relationships or trends. Comparison tests are used to detect significant changes in skill proficiency or application of knowledge at the time of follow-up evaluation (usually 3 to 4 months after training) compared to immediately after training. A relationship or trend analysis may be performed to investigate relationships such as which job-related conditions encourage the graduate to engage in trained behaviors and which ones discourage such behavior.

Summary. The purpose of the follow-up evaluation is to determine the extent to which trained behaviors have transferred to the work environment. This reflects the ultimate value of training to the organization because skills learned in training but not applied on the job are of little value. Instruments used to collect data in the follow-up evaluation are most commonly questionnaires. However, this information may be supplemented with interviews, tests, and performance records. Data analysis usually focuses on testing comparisons and detecting relationships. The results of these analyses reveal significant behavioral changes due to training and identify variables within the work environment which encourage, as well as those which discourage, the graduates to engage in trained behaviors.

Impact Survey

The purpose of the impact survey is to measure effects or impacts of training on the attitudes and socialization skills of the graduates when they return to the workplace. This includes attitudes toward work and socialization skills with co-workers, supervisors, customers, and others. As Ruskin (1919) stated, "The highest reward for toil is not what you get out of it but what you become because of it." In this sense, "what you become" has to do with personal identity and interpersonal relationships within the workplace--in other words, attitudes. The impact survey can furnish valuable information concerning the practical effectiveness of a training program as it is reflected in the graduates' attitudes within the workplace.

Data Collection and Analysis. Data collection for the impact survey could be accomplished by use of instruments such as questionnaires or attitude surveys to be completed by either supervisors, graduates, or both. This information could be supplemented with interviews and review of performance data such as performance appraisals. Data analyses would be primarily comparisons. Usually, this would involve comparing responses from the supervisor to those from the graduate.

Learning Versus Performance. The trainer must be aware of the difference between learning work attitudes and socialization skills versus performing them. Nadler (1990) defines *learning* as "the process of acquiring some new skill, attitude, or knowledge" (p. 110). The training graduate may know how he should perceive and react to his work environment. However, whether he acts on this knowledge when he returns to the work site determines the true effectiveness of the training in improving attitudes and socialization skills of the graduate. This

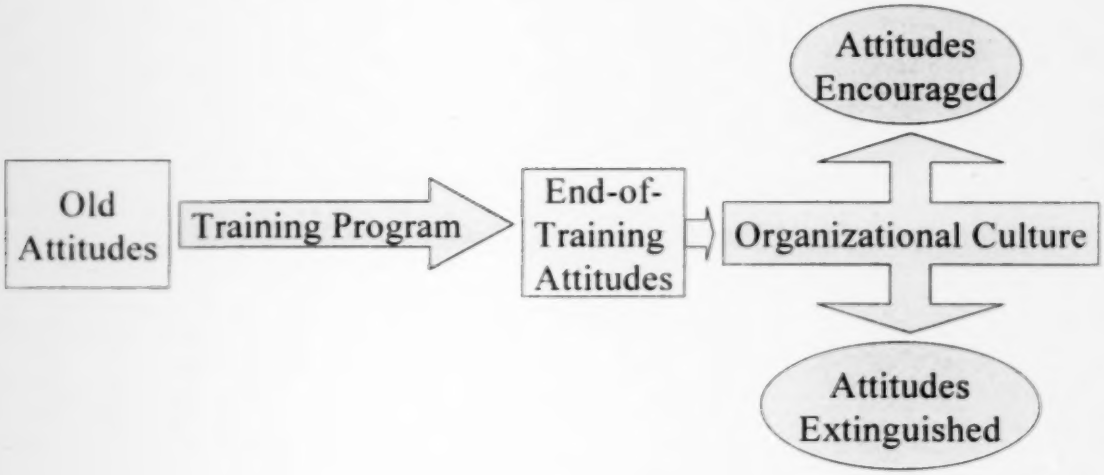
manifestation of the training in action is performance. Nadler (1990) defines performance as "using a skill, attitude or knowledge" (p. 111).

Attitudes — A Tangled Web. While attitudes influence quality of work and quality of work life, the connection between training and attitudes is often difficult to establish. The organizational culture in which the graduate works can quickly extinguish or greatly enhance desirable attitudes, learned and performed during training. If the organizational culture at the workplace supports, or can be modified to support, desirable attitudes, the effectiveness and lasting impact of the training on both the graduate and others within the workplace will be enhanced. The graduate's supervisor has a strong influence on whether desirable attitudes learned in training transfer to the workplace. Phillips (1983) states "What a [supervisor] expects of subordinates and the way he or she treats them largely determines their work performance and career progress" (p. 43). Phillips goes on to explain that creating high expectations is one of the most effective methods a supervisor can use to improve attitudes of subordinates.

Summary. The purpose of the impact survey is to measure the effects of training on the attitudes of training program graduates when they return to the workplace. Such attitudes include socialization skills and the graduates' attitudes toward their work. These attitudes have a strong influence on quality of work performance.

Desirable attitudes learned in training must also be performed within the work setting. However, factors other than the quality of training affect whether learned attitudes are performed within the workplace. One complicating factor is the organizational culture of the employing organization. This is illustrated in Figure 1.

Figure 1. Organizational Culture Modifies Training Effect on Attitudes



Note. Adapted from "Training by Objectives", by G. S. Odiorne, 1970. New York: Macmillan Company, p. 181.

The supervisor has a strong influence on the graduates' attitudes and resulting behavior within the work environment. If the supervisor has positive attitudes and high expectations concerning graduates, their attitudes are likely to be positive and they will be more inclined to transfer trained behaviors and attitudes to the work environment.

Evaluation Strategies

Key questions related to planning evaluation are who, where, and when (Phillips, 1991).

1. Who will conduct the evaluation?

2. Who will supply information? Typical sources are training participants, their supervisors, previous graduates of the training program, customers, and others.

3. Who will provide information about participants, a peer group or subordinates?

4. Who will analyze the data?

5. Who will interpret the data?

6. Who will conduct the follow-up evaluation?

7. Who will decide whether to discontinue or change the process?

The "who" questions should be answered prior to taking any action. In most cases, one person is responsible for everything: collecting data, analyzing it, and reporting to key people. Some questions regarding "where" could include:

1. Where will evaluations take place; in a classroom, on the job, or a balance of the two?

2. Where will the follow-up study be conducted, off or on the job site?

3. Where will observations be made, at the work station or in a simulation of it? The "when" questions relate to the timing of training evaluation.

4. When will training be evaluated? It should be evaluated before, during, and after the training.

5. When will follow-up evaluation be performed? Normally, it should be performed 3 or 4 months after the training.

Two questions regarding assessment of training are asked by Bakken & Bernstein (1982):

1. Has a change actually occurred?

2. Are any changes that have occurred actually the result of training? (p. 44)

The before and after method of measurement listed above can usually answer the first question. However, this may not be as important as knowing that the changes have occurred because of the training.

Using a control group may be impractical or unethical under circumstances which could result in injury to the trainee or others. Under such circumstances, a more feasible approach might be to compare performance of the same group of trainees over time to determine if performance is improving as training progresses.

According to Kirkpatrick (1976), obtaining "proof" that there has been a behavior change can be complicated, time consuming, and expensive. However, this proof may be important if measurement is to be valid and reliable (p. 18-21). Kirkpatrick suggests the following steps:

1. Measure the behavior of control groups before and after training. Use control groups in order to distinguish changes in performance attributable to the training from extraneous influences or conditions.

2. Measure the behavior of control groups over a period of time (successive post-training measures), after waiting long enough to allow participants to practice the skills on the job. Continue the use of control groups to distinguish changes brought about by training.

Bakken & Bernstein (1982) state that there are differences between immediate and long-term changes which occur as a result of training.

> Acquisition of skills or knowledge is usually immediate. We can assess any changes at the completion of training. Changes in productivity, turnover, employee attitudes and the bottom line are unlikely to occur immediately. Assessing the impact of training on these outcomes is necessarily a long-term proposition. (p. 49)

Another method to measure performance is known as the "balanced score card" (Willyerd, 1997, p. 53). Developed by a management consulting firm in Lincoln, MA, the

balanced scorecard helps to counteract weaknesses in traditional financial measures. It achieves this by tracking the key elements of a company's strategy through the use of financial and operational measures. Willyerd (1997) says, "We need to use performance indicators that look to the future" (p. 53).

Similar to Kirkpatrick's model, the balanced score card (Appendix A) looks at four key areas of performance by answering four basic questions (Willyerd, 1997, p. 53).

1. How do we look to our shareholders?
2. How do customers see us?
3. What must we excel at?
4. Can we continue to improve and create value?

Trainers and performance technologists complain that hard-nosed financial measures are too short-sighted to work much of the time. Critical performance measures checked by the balanced score card help to insure that all main points are checked. It uses checks and balances to keep one area from being over emphasized at the expense of another.

This balanced approach can be achieved by maintaining a constant awareness of the vision of the company and working toward that vision. Three steps that help to maintain awareness of the company vision are explained below.

Step One: Define the Organizational Vision. The vision must be defined for the long-term interest of stakeholders. It must also be feasible, focused, and flexible enough to guide decision-making, as well as easy to communicate.

Step Two: Define the Critical Success Factors. Critical success factors are components of the environment, employee behavior, and other factors which will drive performance toward the

BEST COPY AVAILABLE

vision. For example, increasing responsiveness to customers while decreasing costs and maintaining employee satisfaction would be critical success factors.

Step three: Identify specific critical measurements. These should be viewed from four different but complementary perspectives: (a) financial (i.e. net income, return on net assets, cash flow, etc.), (b) customer (i.e. percentage of repeat customers and customer complaints, number of new accounts, etc.), (c) internal (i.e. development time, engineering change notification, scrap and rework, customer response time, etc.), and (d) innovation and learning (i.e. total process time, cycle time, number of new products developed, employee satisfaction, etc.).

By considering all of the above perspectives, the training evaluator can get a more realistic and accurate estimate of the impact a training program has on the organization. This should include not only financial impacts, but also, performance impacts. In order to get a better idea of financial and performance impacts, it is important to chart the measurement using a performance measurement index (Appendix B). The index can provide an indication of performance of training graduates as it relates to organizational goals. It is also good as a "performance tracking tool" that can be used by customers. Selected items of measurement are listed across the top of each column. These items provide for a group's overall performance and include measurements from all four perspectives described above. The column headings (variables reflecting performance) are established by communications with the customer.

Steps to complete the index, in chart form, include the following:

1. Create tentative column headings.

2. With the customer, finalize the column headings.

3. Establish and record the current performance level for each heading.

4. Establish a realistic one-year improvement goal for each heading. Record it at level 10 on the chart.

5. Establish immediate goals, and list them between levels 3 [current] and 10 [desired].

6. Establish lower performance goals. If performance gets worse, for unpredictable reasons, the performance can still be traced. Assign weights to each item. This step may be subjective, but should be a compromise of the four measures, and should total 100.

7. Calculate the value (Appendix C). Record the percent performance values determined at the end of the month in the appropriate column blanks. Next, select the scale in the chart that matches the performance value for the month for the particular column item. Then trace the value across to the score column marked score and record the matching score in the space at the bottom of each column marked score. Take the score value for each column and multiply it times the weight for that column. Record that answer under the column for each heading in the value blank. Values for each column are totaled and written in the correct month blank. This gives a monthly value to compare to, and any month-to-month increase in value is considered an improvement. This tool will allow you to select aligned measures, balance the scorecard and quantitatively track performance improvement results.

Phase 4: Prepare an Evaluation Report

Preparing an evaluation report is an essential step in the program evaluation process. Evaluators sometimes feel that a written report is not necessary when the evaluation findings are favorable. Other times, the evaluation report is simply overlooked in the program evaluation process. However, omitting this phase can jeopardize the credibility and survivability of the

program. "The real measure of an evaluation's success is the extent to which people read, understand and use the findings given to them" (Kenneke, 1987, p. 14). When a written report exists, information can be readily disseminated to those interested in a program.

Preparation of the report involves a clear and well-defined process. First, the data must be interpreted and conclusions drawn. Second, recommendations should be made based on the findings, and then documented in a written report. Finally, an action plan must be developed based on the recommendations.

Interpret the Data and Draw Conclusions

"Interpretation is the act of explaining or telling the meaning of something" (L'Angelle, 1996, p. 242). Once the evaluation is completed, the data must be interpreted. Findings should be reported accurately, fairly, and impartially. When discrepancies between expected and actual training results occur, it is important to explain the differences in the evaluation report. Conclusions reached must be based on discrepancies between expectations and the actual results that are produced by training. Positive conclusions indicate that results exceeded expectations, while negative conclusions indicate that expectations exceeded results. Once the data are interpreted and conclusions are drawn, recommendations must be made that are consistent with the evaluation purpose and objectives.

Make Recommendations

Recommendations are based on the results of the evaluation activities. Managers, trainers, supervisor and committee members may develop the recommendations. However, developing recommendations typically involves a number of individuals who have played key roles in the evaluation process. Therefore, the first step in writing recommendations will likely

be to assemble a group of such individuals and use of the evaluation data since the recommendations will usually result in program modifications. L'Angelle (1996) suggests the following questions be addressed when analyzing conclusions:

1. Is there sufficient evidence to support the conclusions?
2. Do the conclusions reflect the consensus of the group?
3. Are the conclusions compatible with the evaluation objective(s)?
4. Are the statements precise? Are they logical? (p. 246)

When the recommendations have been formulated, the next step is to prepare the written report.

Prepare the Report

Preparing an evaluation report is a straightforward procedure with a given purpose and scope. The report should outline the vocational program, identify who requested the evaluation, reflect the audience, the objectives, and the overall evaluation plan. The data should be presented as simply as possible with only a few visual aids. If corrective action is needed, follow-up is necessary to make certain the changes take place.

Below is a list of guidelines to help assure credibility and completeness of the report.

1. The evaluation purposes and objectives are identified.
2. All evaluation objectives are addressed in detail.
3. Evaluation activities are described — what was done, how it was done and, where appropriate, why it was done.
4. Relevant information and numerical data are presented in narrative text and graphically displayed using tables and figures.
5. The information and data are analyzed and interpreted.

6. Recommendations about the course or program are based on findings and conclusions.

7. Material is presented in an unbiased, factual manner.

8. The writing style is direct and the wording concise.

9. When corrective action is called for, a plan is presented.

10. Instruments used are included in an appendix or as attachments.

Once the evaluation report is completed, the evaluation team will present their findings.

Develop an Action Plan

Activities associated with the action plan must be specified and prioritized. Continuous feedback is necessary throughout the evaluation process; however, it is a key component of the development of the action plan. When objectives are unclear or activities are not specifically assigned to facilitate completion of the action plan, failure of the training program may be the result. Through the evaluation process, discrepancies can be identified, yet failure to take corrective action defeats the purpose of the entire evaluation process.

The importance of developing a plan for corrective action cannot be overemphasized. Program changes or modifications will only be achieved through the process of developing and implementing an action plan. It is not sufficient to identify a problem and its cause. Training personnel must determine what findings are significant and develop detailed action plans to assure that the necessary changes take place.

Action plans detail the items that need to be addressed, who is responsible for the item and the time frame required to implement the action items. Someone in the training organization

should monitor the activities to ensure that the changes are implemented and that the program changes are consistent with the recommendations made in the evaluation report.

By systematically seeking feedback, analyzing it logically, and devising realistic and practical solutions, program improvement can indeed be an intrinsic and ongoing part of the training effort (L'Angelle, 1996). Consequently, the relevance, effectiveness, and cost efficiency of vocational courses and programs can be maintained. The evaluation report should serve as a model to guide the evaluator through this process so that no step will be overlooked.

References

Bakken, D., & Bernstein, A. L. (1982). A Systematic Approach to Evaluation. Training and Development Journal, 36 (8), 44-51.

Campbell, C. P. (1996). Vocational course/program evaluation: A four-phase approach to evaluation. Unpublished manuscript, The University of Tennessee, Knoxville.

Gilley, J., & Eggland, S. (1989). Principles of human resource development. Reading, MA: Addison-Wesley.

Gunn, W. (1987). Client concerns and strategies in evaluation studies. In J. Nowakowski (Ed.), The client perspective on evaluation. San Francisco: Jossey-Bass.

Kenneke, L. (1987, May). A guide for evaluation of technical and vocational education curricula. (Available from the ILO International Centre for Advanced Technical and Vocational Training, Turin, Italy)

Kirkpatrick, D. L. (1976). Determining training needs: Four simple and effective approaches. Training and Development Journal, 31(2), 22-25.

Kirkpatrick, D. L. (1994). Evaluating training programs: The four levels. San Francisco: Berrett-Koehler Publishers.

L'Angelle, D. (1996). An approach to program evaluation. In C. P. Campbell (Ed.), Education and training for work: Vol. 1, Planning programs. Lancaster, PA: Technomic.

Nadler, L. (1990). Every manager's guide to human resource development. (1st ed.). San Francisco: Jossey-Bass.

Odiorne, G. S. (1970). Training by objectives. New York: Macmillan.

Phillips, J. (1983). Handbook of training and evaluation measurement methods. Houston, TX: Gulf Publishing.

Phillips, J. (1991). Handbook of training and evaluation measurement methods. Houston, TX: Gulf Publishing.

Powers, T. (1995). Introduction to management in the hospitality industry (5th ed.). New York: John Wiley & Sons.

Rhoades, J. (1991). Evaluate regular and cooperative vocational programs: Module. Climax Springs, MD: Quality Training Specialists, Inc. (ERIC Document Reproduction Service No. ED 356-337)

Rowntree, D. (1982). A dictionary of education. Totowa, NJ: Barnes & Noble Books.

Ruskin, J. (1919). Sesame and lillies. London: London, Allen, and Unwin.

Vogt, W. P. (1993). Dictionary of statistics and methodology: A nontechnical guide for the social sciences. Newbury Park, CA: SAGE Publications.

Willyerd, K. (1997). Balancing your evaluation act. Training, 34(3), 53-58.

Zenke, R., & Gunkler, J. (1985). Twenty-eight techniques for transforming training into performance. Training, 22(4), 54, 56-63.

Authors' Biographical Information Sheet

James A. Chandler has been involved in the management and administration of hotels, resorts, and restaurants for twenty years. He has earned the hospitality industry credentials as a Certified Hotel Administrator (CHA) and Certified Hospitality Educator (CHE), and he has served on the professional certification faculty of the Educational Institute of the American Hotel and Motel Association since 1992. He is a third-year doctoral trainee and Graduate Teaching Associate at The University of Tennessee, Knoxville.

Robert H. Fulwood is currently enrolled as a second-year masters trainee majoring in Human Resource Development at The University of Tennessee, Knoxville. He previously worked in the U.S. Department of Interior, Office of Surface Mining. In that position, he was involved in developing training related to his professional expertise, revegetation of land disturbed by mining. He was also involved in developing and implementing a training program for Knoxville Field Office Personnel on Total Quality Management (TQM).

Lynn M. Hill has developed and conducted human relations training courses for firms throughout the country, and she has provided a variety of HRD-related consultation services to numerous organizations. Presently, she is a first-year doctoral trainee and Graduate Assistant majoring in Human Resource Development at The University of Tennessee, Knoxville.

Jeanie B. Sharp has spent ten years working in industrial and manufacturing economic development and industrial research and technical development. She is currently a second-year doctoral trainee majoring in Education with a concentration in Leadership Studies at The University of Tennessee, Knoxville.

Keith A. Vincent is currently employed as window distribution clerk at the U.S. Post Office in Madisonville, Tennessee. He served in the U.S. Navy for 20 years in the field of personnel. In that capacity, he was involved in officer and enlisted recruiting, job classification, billet management, education and training, and other duties related to human resource development. He is now enrolled as a second-year masters trainee majoring in Human Resource Development at The University of Tennessee, Knoxville.

Charlie H. Womac is a second-year masters trainee majoring in Human Resource Development at The University of Tennessee, Knoxville. He has fifteen years experience in industry working in quality control and engineering departments. Over the last four years, he has been employed as an instructor for Cleveland State Community College. His primary duties in that capacity has been teaching adults in technical skills.

Appendix A

The Balanced Scoreboard

1. Define the Vision	2. Define the Critical Success Factors	3. Identify the Critical Measurements
To shareholders: Financial Perspective	→	→
To customers : Customer Perspective	→	→
With internal management processes: Internal Perspective	→	→
With companies ability to innovate and grow: Innovation and learning	→	→

Note. Adapted from "Balancing Your Evaluation Act," by K. Willyerd, 1997, Training 34, p. 54.

Appendix B

Performance Measurement Index
Design Group

Percent of Reworked Drawings	Hours Expended per Drawing	Percent Overtime	Department Expense per Drawing	Percent of Drawings Overdue	Percent of Training Time	Dept. 42 Month Jan
						PERFORMANCE
						10
						9
						8
						7
						6
						5 SCORE
						4
						3
						2
						1
						0
						SCORE
						WEIGHT
						VALUE=

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

Note. Adapted from "Balancing Your Evaluation Act," by K. Willyerd, 1997, Training 34, p. 56.

Appendix C

Performance Measurement Index
Design Group

Percent of Reworked Drawings	Hours Expended per Drawing	Percent Overtime	Department Expense per Drawing	Percent of Drawings Overdue	Percent of Training Time	Dept. 42 Month Feb
19%	12.5	6.1%	314	27%	5%	PERFORMANCE
0-5	5.3	0.0	143	15	10	10
6-8	6.5	0.5	164	17	9	9
9-11	7.7	1.0	185	19	8	8
12-14	8.9	1.5	206	21	7	7
15-17	10.1	2.0	227	23	6	6
18-20	11.3	2.5	248	25	5	5 SCORE
21-23	12.5	3.0	269	27	4	4
24-26	13.7	3.5	290	29	3	3
27-28	14.9	4.0	311	31	2	2
29-30	16.1	4.5	332	33	1	1
Over 30	17.3	5.0	353	35	0	0
5	4	0	1	4	5	SCORE
10	30	10	20	20	10	WEIGHT
50	120	0	20	80	50	VALUE = 320

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
300	320										

Note. Adapted from "Balancing Your Evaluation Act." by K. Willyerd, 1997, Training 34, p. 57.

Appendix D

Glossary of Terms

Audience = As used in this report, training audience or audience refers to the decision-makers, trainees, and others for whom the training evaluation was performed.

Comparison tests = Statistical tests designed to determine if there is a significant difference between measures (usually scores on a test) of two or more individuals or groups.

Construct validity = The degree to which a measurement actually reflects the construct being measured. Refer to the definition of "Construct".

Construct = An attitude, feeling, or other theoretical concept which cannot be measured directly. For example, job satisfaction cannot be measured directly, but can only be indirectly assessed from how the employee responds to questions and the evaluator's interpretation of the employee's behavior.

Control group = A group of employees similar to the group that participated in the training program being evaluated, but who did not receive the training. The evaluator strives to select individuals similar to training participants on as many variables as possible. This ensures that differences in the two groups can be reasonably assumed to be due to the effects of the training program.

Curriculum planners = The person or persons responsible for planning the content, methods of presentation, evaluation, and philosophical basis of a training program.

Evaluation = As used in this report, evaluation is a process of assessing the value or benefit of a training program.

External validity = How effectively results of a training evaluation or some component of it can be generalized to evaluations of other similar training programs or to application of the training program under study to other groups.

Extraneous influences = Factors affecting variables the evaluator is measuring, but which are not accounted for in the evaluation and which may have escaped the evaluator's awareness.

Face validity = The extent to which variables selected to measure a construct seem to make sense. Face validity is usually determined by requesting input from experts in the field.

Formative evaluation = Evaluation of a training program while the training is in progress and/or at the close of training. The focus is on how to improve the training for present trainees and for future presentations of the training to other groups.

Gain scores = The numerical difference between scores of training participants on the test taken at the close of training (post-test) minus the score made on the same test at the beginning of training (pre-test). The results represent an estimate of how effectively the training conveyed knowledge, skills, and attitudes it was designed to teach.

Instrument = A device for gathering information about a training program or other subject under study in a form that can be converted to quantitative data for statistical or other analyses. In this report, "instrument" refers only to written instruments such as questionnaires, interview scripts, etc.

Internal validity = The degree to which changes in behavior of the training graduate can be attributed to effects of the training program and not to flawed procedures used in measuring behavior.

Key decision makers = The person or persons with responsibility and authority to make and implement decisions relating to training.

Needs analysis = A process of identifying and prioritizing the needs of a training evaluation audience.

Organizational culture = A prevailing general perception the employees or members of an organization share about the organization. This includes the nature of decisions, the social climate, relationships with others within the organization, and the relationship of the organization with the external environment.

Relationships or trends tests = Statistical tests designed to determine if there is a significant relationship between measures or other quantitative representations of two or more individuals or groups.

Relevance = The degree to which measures used in the evaluation relate to the needs of the evaluation audience.

Reliability = The quality of an instrument that represents the degree to which it will produce the same results if used repeatedly on the same person or group.

Significant = A statistical term meaning a real difference attributable to differences or relationships between two or more groups or individuals and not just a difference due to chance alone.

Socialization skills = Skills in nurturing and maintaining interpersonal relationships and in successfully adapting and functioning within the culture of an organization.

Structured interview = An interview carefully planned in advance with a written protocol or list of questions and instructions specifying how to conduct the interview. Providing structure for the interview helps the interviewer to minimize his or her own influence on the interviewee's responses and enables the interviewer to categorize and quantify responses across all interviewees.

Summative evaluation = A training program evaluation performed after the training is completed in order to get a measure of the overall effect of the training program. The focus is on measuring results of training and not on discovering how to improve the training.

Transfer = The extent to which a training participant applies skills, knowledge, and/or attitudes taught in a training program when he or she returns to the job. Transfer is measured by a follow-up evaluation of the training graduate approximately 3 or 4 months after completion of the training.

Triggering events or situations = This refers to an event or circumstance within the workplace which should cause the graduate of a training program to engage in the behavior learned during training. An example would be the fire fighter who has been trained to quickly put on his fire-fighting clothing and slide down the pole in response to the fire bell. The ringing of the fire bell would be the triggering event for the behavior learned in training--putting on fire-fighting clothes and sliding down the pole.

Validity = The quality of an instrument that represents the degree to which the instrument measures what it is designed to measure and is " . . . free of systematic error" (Vogt, 1993). Validity is classified into internal, external, construct, and face validity. Refer to definitions of these terms to gain a better understanding of validity.

Variable = Any quantitatively measurable factor for which measurement values vary over time.

Vocational training = "Training designed to teach the knowledge, skills, and attitudes required for proficiency in a certain job or task" (Rowntree, 1982, p. 344).

BEST COPY AVAILABLE